ABSTRACT

A computational method system, and computer program are provided for inferring functional links from genome sequences. One method is based on the observation that some pairs of proteins A` and B` have homologs in another organism fused into a single protein chain AB. A trans-genome comparison of sequences can reveal these AB sequences, which are Rosetta Stone sequences because they decipher an interaction between A` and B. Another method compares the genomic sequence of two or more organisms to create a phylogenetic profile for each protein indicating its presence or absence across all the genomes. The profile provides information regarding functional links between different families of proteins. In yet another method a combination of the above two methods is used to predict functional links.

10016819.doc